November 1, 2024

Hon. Ona T. Wang
Daniel Patrick Moynihan
United States Courthouse
500 Pearl St.
New York, NY 10007

Re: *Authors Guild et al. v. OpenAI, Inc., et al.* (No. 1:23-cv-08292-SHS);
        Alter et al. v. OpenAI Inc., et al. (No.1:23-cv-10211-SHS);
        The New York Times Co. v. Microsoft Corp., et al. (No. 1:23-cv-11195-SHS);
        Daily News, LP, et al. v. Microsoft Corp., et al. (No. 1:24-cv-03285-SHS); and
        The Center for Investigative Reporting, Inc. v. OpenAI, Inc., (No. 1:24-cv-04872-SHS)

Dear Judge Wang:

        Following Your Honor's instructions at Wednesday's status conference, I write on behalf of all parties in the above-captioned cases to provide an update on the parties' meet-and-confer efforts related to early Rule 30(b)(6) depositions.

Class & News Plaintiffs' Position: Your Honor recognized that an early 30(b)(6) custodial deposition would "help focus document discovery" and resolve many of the parties' disputes. *See* 10/30/24 Tr. 8:2-20. Plaintiffs have agreed to serve two consolidated deposition notices—one to OpenAI and the other to Microsoft—with each deposition limited to 7 hours.[1] Plaintiffs request that the deposition take place before December 3, 2024. OpenAI has agreed to a custodial deposition so long as certain conditions are met, while Microsoft refuses a custodial 30(b)(6).

OpenAI Agrees To A Deposition, But There Are Three Ripe Disputes For Resolution. OpenAI has accepted that an early custodial deposition is appropriate, subject to the following conditions.

Issue #1 – OpenAI's Custodial Deposition Shouldn't Be Conditioned On Reciprocal Depositions. Class Plaintiffs requested an early Rule 30(b)(6) deposition early in discovery and, until this week, OpenAI has refused. While OpenAI now agrees that an early document 30(b)(6) deposition is appropriate, it refuses to designate a witness unless and until the *Plaintiffs* (News Plaintiffs and the Authors Guild) also put up a witness. That is a delay tactic. OpenAI never once asked for an early deposition from Plaintiffs until yesterday, and has not even sent proposed topics, length, or format. There is no reason to hold Plaintiffs' document 30(b)(6) of OpenAI hostage to a deposition of Plaintiffs on an entirely unspecified subject matter.

Issue #2 – OpenAI's Custodial Deposition Should Occur On Or Before December 3, 2024. Plaintiffs served their topics, the parties will confer, and there are 18 business days until December 3. The deposition should not be delayed; the purpose of an early 30(b)(6) is to alleviate the information asymmetries in this unusually technical and complex case. In a similar copyright class action against Anthropic, Judge Alsup ordered the Defendant to prepare for an early Rule 30(b)(6)

---

[1] Plaintiffs' list of topics is attached as Exhibit 1.

immediately. *See* Ex. 2. That deposition is scheduled for November 8, less than four weeks after the initial case management conference and before any discovery was exchanged.

Issue #3 – *7 Hours Is Appropriate.* Class and News Plaintiffs have already agreed to streamline by taking one deposition of each Defendant, even though each set of cases involves distinct issues. A 7-hour deposition is, if anything, *less* than what Plaintiffs would be entitled to under FRCP 30. *See* Tr. 9:7-8. A more constricted amount of time is inappropriate given that the Class and News Plaintiffs are sharing time. The Class Plaintiffs, for example, will need to address questions about OpenAI retention and deletion of the books1 and books2 datasets—something not at issue in the News cases—while the News Plaintiffs will ask questions about datasets containing news content.

Microsoft's Refusal Of An Early 30(b)(6) Warrants The Court's Intervention. An early 30(b)(6) of Microsoft would advance this case by providing clarity about a variety of issues currently in dispute. Such issues include the identity of custodians with knowledge of the LLMs at issue; Microsoft's searches of its custodians' text messages (none of its productions have included texts so far); and the location of documents and data shared between OpenAI and Microsoft. With respect to the Class Plaintiffs: neither Microsoft's supplemental interrogatory responses nor Microsoft's agreement to produce texts for three specific custodians resolved these disputes. Microsoft has also failed to answer basic questions about what specific models power its products, including on a meet-and-confer last week, making it difficult for Plaintiffs to frame appropriate discovery requests and streamline their case. And Microsoft has been unclear about whether it has searched for and produced documents concerning OpenAI's training data and processes, a critical aspect of this case that all Plaintiffs need to understand—and an issue that is far from resolved, despite Microsoft's claims to the contrary (*see Author Guild* Dkt. 212 at 3-4).

Furthermore, and most critical for the News Plaintiffs, Microsoft does not dispute that it should designate a witness on "output data." It simply contends that now is not the appropriate time for this deposition (without specifying when would be appropriate other than "in due course"), and that such a deposition is not a "custodial" one. Now is exactly the right time, and it makes no difference whether the deposition is labeled "custodial" or something else. As discussed at the October 30 conference, output data–which Microsoft's own counsel described as "actually quite complicated"–is central to various aspects of News Plaintiffs' case. Tr. 54:16-60:11. Plaintiffs need to understand what data has been maintained, how that data is organized, what it contains, and how it is stored so they can formulate an efficient and cost-effective plan for analyzing it with their experts. Delaying the deposition will simply add additional cost, time, and inefficiencies.

Last, even if the Court believes that Microsoft should not be required to put up an early 30(b)(6) witness on certain topics, that does not justify Microsoft's refusal to participate in such a deposition. Plaintiffs are willing to meet and confer with Microsoft on the scope of the topics, but they are not willing to withdraw their request for such an efficient and beneficial deposition.

Microsoft's Position: Rather than identify any ripe issue with Microsoft that an early deposition would help resolve, Plaintiffs are sweeping Microsoft into the disputes they have with OpenAI. ***Microsoft and OpenAI are separate companies.*** Each has its own legal team dealing with different issues arising from Plaintiffs' complaints. Plaintiffs blatantly ignore the Court's clear warning at the hearing that the Court's words should not be used to say "the Judge said we could have a 30(b)(6) deposition," which are all the more applicable to Microsoft. Tr. 8:15-17.

Plaintiffs have not sought *any* relief from the Court related to Microsoft's custodians, sufficiency of document collection, or document retention. In fact, at the last hearing, Plaintiffs did not disagree with Microsoft's representation that there were no open custodial disputes against it. Yet nearly all of 18 custodial deposition topics sent to Microsoft seek exactly that. Discovery on discovery is not a casual tool to be used for case management – it is "disfavored" and therefore requires a significant showing that Plaintiffs fail to even begin to provide here. *See Orillaneda v. French Culinary Inst.*, No. 07-cv-3206, 2011 U.S. Dist. LEXIS 105793, at *24 (S.D.N.Y. Sept. 19, 2011) ("[A] plaintiff is not entitled to conduct discovery that is solely relevant to the sufficiency of the adversary's document production without first identifying facts suggesting that the production is deficient."); *Trustees of Local 854 Pension Fund v. Barrett*, No. 23 Civ. 1160, 2024 WL 4026257, at *5 (S.D.N.Y. Sept. 3, 2024). And the vast breadth of the topics is not tailored to facilitate efficient completion of discovery—which was the Court's guidance. Tr. 8:2-20.

Moreover, Plaintiffs misrepresent the state of affairs vis-à-vis Microsoft. Plaintiffs' first example (individuals knowledgeable regarding LLMs) was addressed with a supplemental interrogatory response served earlier this week. Their second (text messages) was not raised in any meet and confer. Regardless, Microsoft will produce text messages it agreed to produce to the Class Plaintiffs (represented by the same law firm). Their third example (what models are used by what products), is a merits issue not a custodial issue, and an interrogatory would be far more efficient. Finally, with respect to the question of Microsoft's search and production regarding OpenAI's training data, Microsoft served supplemental responses after the September 12 Status Conference *that resolved the issue*. *See* Dkt. 212 at 4-5. Class Plaintiffs notably *did not* include it (or any of these other issues) as an open issue on the agenda for the October 30 status conference. Dkt. 220.

The News Plaintiffs' request for a deposition relating to output data is plainly not a "custodial" deposition. The 15 petabytes of prompt data is not in the possession of any individual custodian. Moreover, no deposition is needed about *where* relevant documents might be found because Plaintiffs *already know where*. This deposition is not designed to facilitate gathering evidence for the merits. The News Plaintiffs concede it is for their experts. The sheer size of the data set and other issues related to it creates a major burden to conduct a deposition on short notice. A deposition on this topic will happen in due course, but there is no reason to upend the normal case schedule and rush into it now.

OpenAI's Position: OpenAI has conferred with Plaintiffs on the "proper scope" for "limited 30(b)(6)" depositions, for both sides, on "custodial" topics to "help focus the document discovery." Dkt. 217; 10/30/24 Tr. at 8:2-5, 63:3-4. That conferral has revealed that Plaintiffs seek a deposition on improper topics, while insulating themselves from similar discovery.

As an initial matter, OpenAI has not required any conditions on reciprocity. The Court indicated that Defendants may serve "a cross 30(b)(6) for plaintiffs" to address "the early issues that seem to be holding up some of the discovery." *Id.* at 8:13-15. Defendants will do so. The 30 Authors Guild Plaintiffs have produced fewer than 14,000 documents combined. 75% of the NYT's production consists of copyright registrations, and agreements with reporters. Beyond deposit copies of their works, the 8 Daily News Plaintiffs have produced fewer than 1,500 documents combined. Most Plaintiffs also have not produced text messages from their own custodians.

3

The scope and timing of Plaintiffs' proposed early 30(b)(6) of OpenAI is also improper. *First,* Plaintiffs' topics seek privileged testimony regarding details about Defendants' attorney-led investigations, document preservation notices, and retention policies. (Topics 2, 3 & 7). These topics address privileged and work product protected information and are thus improper. *See Robert W. Gordon, Esq. v. City of New York*, 2016 WL 3866585, at *2 (S.D.N.Y. July 13, 2016).

Second, Topics 10-18 seek premature discovery into merits issues concerning Defendants' generative AI products and services, text generation AI models, output logs, and more. None of this is proper custodial discovery and now is not the time for merits depositions.

Third, it makes little sense for the parties to take custodial depositions in the next several weeks, when significant document discovery on both sides remains that will narrow the custodial issues. Plaintiffs cite the Anthropic case, but the issue there had nothing to do with custodial depositions. The court authorized a narrow deposition *or* written discovery limited to "whether plaintiffs' works are/were in the defendant's training set." *Bartz v. Anthropic*, Case No. 24-cv-5417, Dkt. 46 (N.D. Cal.). That limited scope supports OpenAI's position, not Plaintiffs'.

OpenAI will continue to confer, but Plaintiffs' proposal makes clear that guardrails are necessary. The Court should order that for all early custodial depositions: (1) the topics be cabined to truly custodial issues; (2) they should take place in January 2025, after the parties have completed further document discovery; (3) the parties should be limited to 7 custodial topics per side; and (4) each deposition should be limited to 3 hours.

Sincerely,

SUSMAN GODFREY L.L.P.

/s/ Rohit Nath
Rohit Nath